

1	303.4	21.8	1711	11	AKO07734	Mus muscu
2	294.6	21.2	1774	11	BC027097	Mus muscu
3	259.4	18.7	624	10	BZ000527	BB300527
4	247	17.8	810	12	BG675952	BG675952
5	240.6	17.3	676	9	AL652777	AL652777
6	238.4	17.2	392	10	AEO31611	AB031611
						130228 MA

Qy	544	AAACGTATGTCACAGATCCAACGCCCTTCTGGGCTCAGTGCCCCCACTGACCTATCAGTGC 603 Db						
Dd	181	AAACGTATGTCACAGATTCAAGCCCTTCTGGGCTCAGTGCCCCCACTGACCTATCAGTGA 240 						
Qy	604	TATGTGAGCGCTGATCTCTGCAGAGGTTCCCCTGGGTAGAGTTGTG 648 						
Dd	241	TATGTGAGTCTGATCTCAGCAGAAGTCCCCTGGGTAGAGGTGAG 285 						
RESULT 4								
BG675952								
LOCUS		BG675952 810 bp mRNA linear EST 01-MAY-2001						
DEFINITION		602622310P1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4747323 5'; mRNA sequence.						
ACCESSION		BG675952						
VERSION		BG675952.1 GI:13907348						
KEYWORDS		EST.						
SOURCE		human.						
ORGANISM		Homo sapiens						
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.						
AUTHORS		1 (bases 1 to 810)						
TITLE		NIH-MGC http://mgc.nci.nih.gov/.						
JOURNAL		National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)						
COMMENT		Contact: Robert Strausberg, Ph.D. Email: rcgapes@remail.nih.gov Tissue Procurement: James Cleaver, M.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E Consortium (LLNL) DNA sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E Consortium/LLNL at: http://image.llnl.gov Plate: LHAM10597 row: C column: 04 High quality sequence start: 4 High quality sequence stop: 797.						
FEATURES		Location/Qualifiers						
source		1..810 /organism="Homo sapiens" /db_xref="taxon:9606" /clones="IMAGE:4747323" /clone_lib="NCI_CGAP_Skn4" /tissue_type="squamous cell carcinoma" /lab_hosts="DH10B (TI phage-resistant)" /note="Organ: skin; Vector: pCMV-SPORT6; Site: 1: Notf; Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.5Kb. Library constructed by Life Technologies. Note: this is a NCI_CGAP Library."						
BASE COUNT		214 A 178 C 169 G 249 T						
ORIGIN								
Query Match		17.8%; Score 247; DB 12; Length 810;						
Best Local Similarity		99.6%; Pred. No. 1.5e-56;						
Matches		258; Conservative 0; Mismatches 0; Indels 1; Gaps 1;						
Qy	1131	CTTGCGTTTTTAAGTCTTTGGAGTGAAGTGTTACTGATTAATGCTCATTCCTTGATTCG 1190 						
Dd	1	CTTGCGTTTTTAAGTCTTTGGAGTGAAGTGTTACTGATTAATGCTCATTCCTTGATTCG 60 						
Qy	1191	CTTGCGCTCAATAATGCTTATCTGATTTCCATTTGGCTTCAATGCTCTTTTCTTCCAGTA 1250 						
Dd	61	CTTGCGCTCAATAATGCTTATCTGATTTCCATTTGGCTTCAATGCTCTTTTCTTCCAGTA 120 						
Qy	1251	CTTGCACTCAATTGCGCTCACTCTTACCACCAAATGATAGTAGACTACCTCCATTGTGCTG 1310 						
Dd	121	CTTGCACTCAATTGCGCTCACCTCTTACCACCAAATGATAGTAGACTACCTCCATTGTGCTG 180 						
Qy	1311	CTGTFCACGAGACCCTCGGACCGAGGTGAGAACTCAGAGCCACCCTTTGAGACTGAAGC 1370 						
Dd	181	CTGTFCACGAGACCCTCGGACCA - GGTTGAGAACTCAGAGCCACCCTTTGAGACTGAAGC 239 						

```

Db 306 AACTTTAGCTGGTTGGTACTGTGTGACAGCTCCCTGGGTGAGTAACCTTGTCTTTATTCAGCC 365
QY 1003 ATCAACTTCTCTGCTGCTCAGCTTTGAGTTGAGGTTGGCAGACAGAGATCTCGTCGAC 1062
Db 366 ATCAATATGTTTGTCTGCTGCTGCTTTCAGCTAAACCTATCAGATCGGAGCTTAATTCAC 425
QY 1063 AAAGGGCAGAACTGGGGACATATGCGGCTGCACATATAGTGTGAGGTTGGTAGAGAAATGTG 1122
Db 426 CAGTCACAAAACCTGGGGAGGTTAAGTCTTCATTATACCAATTCGGTTGCTCGAGAATGCA 485
QY 1123 ATCATGCTCTGGTTTAAAGTTCTTTGGAGTGAAGTTGTTACTGAATTAAGTCTCAATCC 1182
Db 486 GCCCTCATCTGCTCTGCTGCTTCTACAGGAGGATGTGTTTGAATATTTCTGTTCCTCCA 545
QY 1183 TTGATTGCTTGCAGCTCAATATTCTTATCTGATTTCCATTGGCTTCATGCTCTCTTTTC 1242
Db 546 CTGCTGGTGGTTCAACTTTAGTAGATATGCCACCGCATCTTCTTCATGCTGCTTTT 605
QY 1243 TTCAGTACTTGCATCCATTCGGCTCACTCTTACCCCAATATGATAGACTACCTCCAT 1302
Db 606 TACCAGTACCTTCACTTCCCTTCCGCTTCTCTTTCAGCCCAAGTTGAAGACTGTTTGGCT 665
QY 1303 TGTGCTGCTG 1313
Db 666 TGTGTTGCTG 676

RESULT 6
LOCUS BE031611 392 bp mRNA linear EST 09-JUL-2000
DEFINITION 130228 MARC lPIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BE031611
VERSION BE031611.1 GI:8326620
KEYWORDS EST.
SOURCE Sus scrofa
ORGANISM Sus scrofa
REFERENCE 1 (bases 1 to 392);
AUTHORS Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
TITLE Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine
JOURNAL Unpublished (2000)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCTCAGTCAGG
Plate: 59 row: C column: 17
Seq primer: ATTTAGGTGACATATAG.
FEATURES
source
1. 392
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC lPIG"
/tissue_type="pooled"
/lab_host="DH10B"
/notes="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI; Library made from pooled tissue from day 11, 13, 15, 20, and 30 embryos."
BASE COUNT 81 a 87 c 100 g 124 t
Query Match 17.2%; Score 238.4; DB 10; Length 392;

Db 306 AACTTTAGCTGGTTGGTACTGTGTGACAGCTCCCTGGGTGAGTAACCTTGTCTTTATTCAGCC 365
QY 1003 ATCAACTTCTCTGCTGCTCAGCTTTGAGTTGAGGTTGGCAGACAGAGATCTCGTCGAC 1062
Db 366 ATCAATATGTTTGTCTGCTGCTGCTTTCAGCTAAACCTATCAGATCGGAGCTTAATTCAC 425
QY 1063 AAAGGGCAGAACTGGGGACATATGCGGCTGCACATATAGTGTGAGGTTGGTAGAGAAATGTG 1122
Db 426 CAGTCACAAAACCTGGGGAGGTTAAGTCTTCATTATACCAATTCGGTTGCTCGAGAATGCA 485
QY 1123 ATCATGCTCTGGTTTAAAGTTCTTTGGAGTGAAGTTGTTACTGAATTAAGTCTCAATCC 1182
Db 486 GCCCTCATCTGCTCTGCTGCTTCTACAGGAGGATGTGTTTGAATATTTCTGTTCCTCCA 545
QY 1183 TTGATTGCTTGCAGCTCAATATTCTTATCTGATTTCCATTGGCTTCATGCTCTCTTTTC 1242
Db 546 CTGCTGGTGGTTCAACTTTAGTAGATATGCCACCGCATCTTCTTCATGCTGCTTTT 605
QY 1243 TTCAGTACTTGCATCCATTCGGCTCACTCTTACCCCAATATGATAGACTACCTCCAT 1302
Db 606 TACCAGTACCTTCACTTCCCTTCCGCTTCTCTTTCAGCCCAAGTTGAAGACTGTTTGGCT 665
QY 1303 TGTGCTGCTG 1313
Db 666 TGTGTTGCTG 676

RESULT 7
LOCUS BE031611 881 bp mRNA linear EST 04-OCT-2001
DEFINITION 603075458F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5167070 5', mRNA sequence.
ACCESSION BE031611
VERSION BE031611.1 GI:15939332
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 881)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLAM11415 Row: d column: 15
High quality sequence stop: 782.
FEATURES
source
1. 881
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5167070"
/clone_lib="NIH_MGC_119"
/tissue_type="medulla"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source normal medulla from anonymous male age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.3 kb, insert size range 0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH MGC Library."
BASE COUNT 257 a 163 c 194 g 266 t
Query Match 14.8%; Score 206.2; DB 13; Length 881;
Best Local Similarity 56.8%; Pred. No. 2.5e-45;

```

Matches	484;	Conservative	0;	Mismatches	349;	Indels	19;	Gaps	5;
Qy	8	CAAGACCACAA	CATT	CAGAAAGAA	CTCGACAAT	GGACAGAGTTT	TATGAAAT	TCCTGAGG	67
Db	34	CAAGGAAGCA	AAAAA	ATAAGCA	CCCTGAGAT	TGGACAGTGT	TGAGAGAT	TGGATG	92
Qy	68	AGCCAAATGT	TGGAT	CCGGTTT	CATCTCTG	GAGGAAGAT	TGTCAGT	GGAGCCAA	127
Db	93	AAGAAGCAC	CAGGAGG	TTTTCAT	CTTCGAA	AGAGAAAT	TAGTCTCTGG	CAGAGACTCC	151
Qy	128	GATTTACTTT	TTCCATT	TAGCAT	CTTTTCTCC	ACCTTTT	TGTACT	TGCGGAGGCTGCAT	187
Db	152	ATCTAAGCT	TTTCTTT	TAGCAT	TATCTTCT	CAACTGTTCT	CTACT	TGTGAGGTGTGCT	211
Qy	188	CTGCTTTGT	TACAT	TGGTTAG	AATCTAT	CGAAAGAA	TAGTGA	AACTTACTTGGATGACATACA	247
Db	212	TGCGTTT	TATACAT	GTTTG	AAATTTAT	CGAAAGCT	TAATGAC	NATCTTGGATGTCAATTA	271
Qy	248	CTTTTCTTT	TTTAT	TGTTTTC	ATTCAT	TATATG	GTTCAGT	TGACCTCAATTTTGTCCACA	307
Db	272	CCATCAGCT	TTTAT	TGTGGGG	CAATTTT	TGATCA	AAATAT	CTCTGATGTTTTTCAACA	331
Qy	308	GAGACTAG	CCAAAGAT	AAACCG	CTATCAT	TATATTT	TATGCAT	CTAATCTCTTGGGACCTG	367
Db	332	AAGACTT	GAGGAGAA	T	AGGCTG	CAATTA	CTTTT	TGTGGCACATTTCTCTTTTAGGACCTA	391
Qy	368	TTATCAGAT	TTTTG	GAGGCCAT	GTATTA	AGTACCT	CACACT	GTGGAAGAAAGGAGCAGG	427
Db	392	TNGTGAGGT	TTTTG	CACACCAT	TAGAAAT	TACCA	AAATGGT	TGAAAAATCTTTAAACAGG	451
Qy	428	AGGAGCCCT	TATGT	TCAGCCT	CA	CCCCGAA	AGAAAGAT	GCTAATAGATGGCGAGGAGTGCTGA	487
Db	452	AGAAGG	AGAGACT	CAAGT	TAGCAT	CA	AAAGAG	-----AAACGAGATGC	496
Qy	488	TAGAAAT	TGGAGGT	TGGGCC	CACTCC	ATCCGGA	CCCTGG	CTATGACGCGCAATGCCTACAAAC	547
Db	497	TGGAAG	GGGAGAT	TTCAT	TCTCAAT	CCGGGAT	TAATTT	CAATGAGCAGAAAGCGTTTCAAGT	556
Qy	548	GTATGT	CACAGAT	CCAAGC	CTTCT	CTGGGCT	CAGTGC	CCCGACGTGACCTATCAGCTCTATG	607
Db	557	ACATGT	CAGT	GAATC	AGGCTTT	CTCGGT	TCGTGTT	CCCAATTAATTTTGAGATGTATA	616
Qy	608	TGAGCCT	GTATCT	CTG	CAGAGGT	TTCCCT	TGGGTAG	ATGTTGTGCTAATGGTATTTTCCCTGG	667
Db	617	TCAGTCT	CAC	TATAC	GAGAA	TGGCCTTT	TGAAT	TAGAG-CATGCTGATGACATGTTCCCTGT	675
Qy	668	TATCTGT	CACCTAT	TGGGG	CCACCTTT	TGCAAT	TATGTT	TGGCTATCCAGATCAAGTACGATG	727
Db	676	TATCAGT	TACTT	TATG	GGGCCA	TTCCGCT	GCAATAT	ATACGTGGCATCCAGATCAGCAATGATG	735
Qy	728	ACTACA	AGATTC	CGCCT	TGCGGCC	CACTAG	AAAGT	CTCTCT-CATCACCATCTGGCGGACATTG	786
Db	736	ATATCA	CCATAT	TAAGCT	ACCGGAT	TAGAAAT	TTCTT	CTGTGTGCTGATGGGCGCATTTATGG	795
Qy	787	GAGAT	CACTT	CCCCG	CCTCCT	GTATCT	TGCTGT	GTCTTCTCAGCGACCTTTTGAAATTTGAAGCT	846
Db	796	AAGTTAT	CTC	ACGTG	TAGT	GACTG	GCACTT	TCCCTTGGCATCTCTGAACTTTGTAGAC	855
Qy	847	GTG	CCCTT	CTCCTA					858
Db	856	TAA	CCGTT	TGGTA					867

RESULT 8	
AA256009	
LOCUS	
DEFINITION	201 bp mRNA linear EST 15-AUG-1997 z28c06.r1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:686506 5',
ACCESSION	AA256009
VERSION	AA256009.1
KEYWORDS	GI:1891549
SOURCE	EST.
ORGANISM	human.
	Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 201)
NCI-CCAG <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert length: 1241 Std Error: 0.00
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 183.
Location/Qualifiers
1. .201
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:686506"
/clone_lib="NCI_CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
/note="vector: p773D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGAGCGCGCTCAATTTTTTTTTTTT-3',
]. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p773 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
39 a 46 c 36 g 80 t
BASE COUNT
ORIGIN

Query Match	14.4%	Score 199.4;	DB 9;	Length 201;
Best Local Similarity	99.5%;	Pred. No. 1e-43;		
Matches 200; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;
Qy	1116	GAATGTCATCATGTCTTGTTTTAAAGTCTTTGGAGTGAAGTGTTACTGAATTACTG	1175	
Dd	1	GAATGTCATCATGTCTTGTTTTAAGTCTTAGTGAGTGAAGTGTTACTGAATTACTG	60	
Qy	1176	TCAATCCTTGATGCCCTTGAGCTCATTATTGCTTAATCTGATTTCCATTTGCCTTCATGCT	1235	
Dd	61	TCATTCCTTGATGCCCTTGAGCTCATTATTGCTTAATCTGATTTCCATTTGCCTTCATGCT	120	
Qy	1236	CCTTTTCTTCAGPACHTTGATCCATTGCGCTCACTCTTCACCACAATAATCTAGTAGACTA	1295	
Dd	121	CCTTTTCTTCAGPACHTTGATCCATTGCGCTCACTCTTCACCACAANTAGTAGACTA	180	
Qy	1296	CCTCCATTGTGCTGCTGTCA	1316	
Dd	181	CCTCCATTGTGCTGCTGTCA	201	

RESULT 9	
AL638533/c	
LOCUS	678 bp mRNA linear EST 12-DEC-2001
DEFINITION	AL638533 XGC-neurula <i>Silurana tropicalis</i> cDNA clone TNeu018n12 5' ; mRNA
ACCESSION	AL638533
VERSION	AL638533.1
KEYWORDS	EST. GI:16790512
SOURCE	western clawed frog.
ORGANISM	<i>Silurana tropicalis</i>
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
	Xenopodinae; <i>Silurana</i> .
REFERENCE	1 (bases 1 to 678)

REFERENCE
1 (bases 1 to 678)

AUTHORS Huckle, E., Taylor, R., Ashurst, J. L., Zorn, A. M. and Rogers, J.
TITLE Sanger Xenopus tropicalis EST project 2001 (10_2001)
JOURNAL Unpublished (2001)
COMMENT Contact: Huckle E
Sanger Centre
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS SEQUENCE ID: TNeu018n12.sp6
Sequencing primer: SP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
FEATURES
 source
 1..678
 Location/Qualifiers
 /organism="Silurana tropicalis"
 /db_xref="taxon:8364"
 /clones="TNeu018n12"
 /dev_stage="neurula"
 /lab_host="XGC-neurula"
 /note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
 was oligo dT primed from 5ug of poly A+ RNA from neurula.
 EcoRI-NotI cut cDNA was then ligated into pCS107 with
 EcoRI at the 5' end and NotI at the 3' end."
BASE COUNT 225 a 156 c 153 g 144 t
ORIGIN
Query Match 13.9%; Score 193.4; DB 9; Length 678;
Best Local Similarity 58.7%; Pred. No. 7.4e-42;
Matches 335; Conservative 0; Mismatches 236; Indels 0; Gaps 0;
QY 642 AGTGTGCTAATGATTTTCCTCGTATCTGTCACTATGGGCGCCACCTTTGCAATAT 701
Db 616 AGCTGTCTCATGGCTATGTGCTTCCTCAACTACATATGGTGCCTTAATCCTAAATG 557
QY 702 GTTGGCTATCCAGATCAAGTACGATGACTACAAAGATTCGCTTGGGCCACTAGAAGTCT 761
Db 556 CTTGGCCATACAAATTAAGTATGATGACTATAAAGTTGCTAAGTGGCCAGGTTCT 497
QY 762 CTGATACCATCTGGGACATTTGGAGTACATCTCCGCCCTCTGATTTCTGGTCTTT 821
Db 496 GTGATCTGTGATGGAGATGCTGGAGATGCTACCCGTGTTACAGATTTCTGGTCTTT 437
QY 822 CTCAGCCACTTTGAATTAAGCTGTGCTTCTCTAGTCTCAACTTCTGATCATCTCT 881
Db 436 TTGTTACGCTCTTAAGCCCTGGGTGCTGCTGAGTGGCTTAGCAAACTTACTGGCGCTCT 377
QY 882 CTTTGAGCCCTGGATTAAGTCTGGAGAAAGTGTGCCAGATGCCCAATAACATTGAGAA 941
Db 376 CTTTGTGCTTGGTTAAATTTCTGGAGCAGTGGGACAGCTGGCCCAAAACATAGAAA 317
QY 942 AAATTTAGCCGGTGGGCACTCTGTGTGCTCTGATTTTCAATCACCCTCTATGCTGG 1001
Db 316 GAACTTTAGCTGTTTGGTGTCTGTGACAGTCTCTGGTTCAGTAACTTCTTTATTCAG 257
QY 1002 CATCAACTCTCTGCTGCTGCTTGTGAGTTGAGTTGGGAGCAGAGATCTGTCGA 1061
Db 256 CATCAATATGTTTGTGCTGCTGCTTCAAGCTTAACTTATCAGATCGGAGCTTAATGA 197
QY 1062 CAAAGGCGAGAACTGGGACATATGGGCTGCACTATAGTGTGAGTTGGTAGAGATGT 1121
Db 196 CCAGTCACAAACTGGGAGAGTTAAGTCTTATATACCATTCGTTGCTCGAGATGC 137
QY 1122 GATCATGCTTTGGTTTTTAAAGTCTTTTGGAGTGAAGTTGTTACTGAATTTACTGTCATTC 1181
Db 136 AGCCCTCATCTGCTGTGCTACTTCTACAAGGAGGATGTGTTTGAGTATTTCTGTTCGCC 77
QY 1182 CTTGATTCGCTTGCAGCTCATTTATGCTTAT 1212
Db 76 ACTGCTGGTGTTCAACTCTTAGTAGGATAT 46
RESULT 10

BI546216
LOCUS 60318602P1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:5259961.5',
DEFINITION mRNA sequence.
ACCESSION BI546216
VERSION BI546216.1 GI:15433528
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 793)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM11655 row: c column: 02
High quality sequence stop: 707.
FEATURES
 source
 1..793
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clones="IMAGE:5259961"
 /clone_lib="NIH_MGC_95"
 /tissue_type="hippocampus"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pBluescriptR (modified
 pBluescript KG+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTGTGTTTTTTTNN-3',
 size-selected for average insert size 2.5 kb and
 normalized to ROT 5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NIH/NHGRI, National
 Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 180 a 215 c 188 g 210 t
ORIGIN
Query Match 13.1%; Score 182.6; DB 13; Length 793;
Best Local Similarity 56.9%; Pred. No. 7.2e-39;
Matches 419; Conservative 0; Mismatches 304; Indels 14; Gaps 4;
QY 289 ACCCTCATTTTGTCCACAGAGATCTAGCCAAAGATAAACCGCTATCATTTATGTCAT 348
Db 1 AGCGGCTCTTCTGTCACCGGACCTCAGCCGCGACCGCCGCTCGTACTGCTGCTGCAC 60
QY 349 CTAATCCTCTTGGGACCTGTTATCAGATGTTTGGAGGCCATGATTAAGTACCTCACACTG 408
Db 61 CTGCTGCAACTTGGGCGCGTTTTCAGGTGTTTGAAGTCTTCTGTCATCTAC----- 111
QY 409 TGGAAAGAGAGGAGGAGGAGGAGGCGCTTATGTGAGCTCACCCGAAAGAG---ATGCTA 465
Db 112 TTTTCAGTCAGGCAACAACTGAAGAGCCTTATGTCACTATCACCAGAGAGAGGCAATGCCA 171
QY 466 ATAGATGGCGAGAGGTGCTGATAGAATGGGAGGTGGGCCACTTCATCCGAGACCTGGCT 525
Db 172 AAAAATGGCTCTCAGAGAGATTGGAAGAGGTGGGCCAGGAGCAAGCAACTAATC 231
QY 526 ATGACCCGCAATGCCCTACAAAGTATGTACAGATCCAAAGCCTTCTCTGGCTCAGTGCCC 585
Db 232 ACCCACCATCATCGCTTCAGCCGGCGTGGTATCCAGGCTTCTTGGGCTCAGGCCCC 291
QY 586 CAGCTGACCTATCAGCTCTATGTGAGCCTGATCTCTCTGAGAGGTTCCCTCTGGGTAGAGTT 645

Db 292 CAGCTGACCCCTACAGCTGTACATTAAGTGTGTCATGACGAGAGAGTCACTGTGTGGAAGAAGT 351
 QY 646 GTGCTAAATGGTATTTTCCCTGTATCTGTACCTATATGGGGCCACCTTTGCAATATGTTG 705
 Db 352 CTCTCATGACCATATCCCTGTGTGTCATTTGTGTATGGAGCCCTTGGCGTGCACATCTTA 411
 QY 706 GCTATCCAGATCAAGTACGATGACTAC-AAGATTCGCCCTTGGGCCACATAGAAAGTCTCTG 764
 Db 412 GCCATCAAAATCAAGTACGATGAGTATGAAGTCAAGAGTGAAGGCTCTGGCCATGTCTG 471
 QY 765 CATCACCATCTGGGGGACATTTGGAGATCACTTCGCCCTCTCTGATCTGGTCTCTCTC 824
 Db 472 TATCTTCTGCTGGAGGAGCTTTGAGATTGCCACTCGAGTTGTAGTCTGCTCTCTTTAC 531
 QY 825 AGCCACTTTGAAATGAAGGCTGTGCCCTTCTCTAGTGTCTCACTTCCCTGATCATCTCTT 884
 Db 532 CTCCGCTCTGAAGACCTGGCGGTGGTTATATATCTCATCACTTCTTTCAGTTCTTCTTA 591
 QY 885 TGAGCCCTGGATTAAGTTCCTGGAGAGTGGTCCAGATGCCCAATAACATTTAGAAAAA 944
 Db 592 GTACCCCTGGATCTCTTCTGTGTGAGTGGTTCCTCCATTCCTCGAGACATAGAG-AAGG 650
 QY 945 CTTGAGCCGGTGGGACATCTGGTGGTCTGATTTTCAGTCACCATCTCTATGCTGGCAT 1004
 Db 651 CCTCAGTAGAGTGGGACCAACCATTTGATGATGATCTTAACTTTTACTCTATGCTGGTA 710
 QY 1005 CAACCTCTCTGCTGGT 1021
 Db 711 TTAACCATGTTCTCTGCT 727

RESULT 11
 B1463670
 LOCUS 603207340F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5273312 5',
 DEFINITION mRNA sequence.
 ACCESSION B1463670
 VERSION B1463670.1 GI:15254326
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 855)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L14M11689 row: 0 column: 09
 High quality sequence stop: 826.
 Location/Qualifiers
 1..855
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5273312"
 /lab_host="DH10B"
 /note="Organ: testis; Vector: pBluescriptR (modified
 pBluescript KS+); Site:1: BamHI; Site 2: SalI-XhoI (gtcgag
); Oligo-dr primed using primer 5'-TTTTTTTTTTTTTTVN-3',
 size-selected for average insert size 2.2 kb and
 normalized to ROT 5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library

constructed by M. Brownstein (NIMH/NHGRI, National
 Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 257 a 156 c 186 g 256 t
 ORIGIN

Query Match 12.5%; Score 174.2; DB 13; Length 855;
 Best Local Similarity 57.0%; Pred. No. 1.5e-36;
 Matches 347; Conservative 0; Mismatches 248; Indels 14; Gaps 1;

QY 8 CAGAGCCCAACATTCAGAAAGAACTCGACATGAGAGAGTTTATGAAATTCCTGAGG 67
 Db 72 CAAGGAAGCGAGAAAAATAAGCACCCCTGAGAAATGAGACAGTGTGTTGAAGAGATGGATG 131
 QY 68 AGCCAAATGTGATCGCGTTTCATCTCTGAGAGAGATGTCTCCGTGGAGCCAAACCC 127
 Db 132 AGAAAGCACAGGAGAGTTTCATCTTGAAGAGAGAAATAGTCTTGGCCAGAGACTCC 191
 QY 128 GATTATCTTTTCCATTAGCATCTCTTTCACACCTTTTTCACCTTTTGTACTGTGGGAGGCTGCAT 187
 Db 192 ATCTAAGCTTTCTCTTTTAGCATTTCTTCACTTGTCTTCTACTGTGGTGAAGTTGCCT 251
 QY 198 CTGCTTTGTACATGTTAGAAATCTATCGAAAGAAATAGTGAAGAACTTACTGGATGACATACA 247
 Db 252 TTGGTTTATACATGTTTGAATTTTATCGAAAGAACTTAATGACACATTTCTGGATGTCATTTA 311
 QY 248 CCTTTTCTTTCTTTATGTTTTCATCCATTTATGTCCTGTCAGTTGACCCCTCATTTTGTCCACA 307
 Db 312 CCATCAGCTTTTATTTGTTGGGGCAATTTTGGATCAAAATATCTCTGATGTTTTCACA 371
 QY 308 GAGATCTAGCCAAAGATAAACCGCTTATCATTTATTTATGATCTTAATCTCTTGGGACCTG 367
 Db 372 AAGACTTGAGGAGAAATAAGGCTGCATTTCTTTTGGCATTCTTCTTTTAGGACCTA 431
 QY 368 TTATCAGATGTTTGAGGCCATGATTAAGTACTCCTACACTGTGGAAGAAAGAGGAGGAGG 427
 Db 432 TTGTGAGGTGTTTGGCACACCATTAGAAATTTACCAAAATGGTTGAAAAATCTTTAAACAGG 491
 QY 428 AGGAGCCCTTGTCTGACGCTCACCCCGAAGAGATGCTATAGATGCGGAGGAGGTGCTGA 487
 Db 492 AGAAGGAAGA-----GACTCAAGTTAGCATCAAAAGAGAAACACGATGCT 537
 QY 488 TAGAATGGAGTGGGCCACATCCATCCGACCCCTGGCTATGACACCCCAATGCTCAAAAC 547
 Db 538 GGCAAGGGAGATTGCATTTCAATCCGGGATAATTTTCATGACAGAGAGGCTTTCAAGT 597
 QY 548 GTATCTCAGATCCAAAGCCTTCTGGGCTCAGTCCCGAGCTGACCTATCAGCTCTATG 607
 Db 598 ACATGTCAGTGATTCAGGCTTTCTCGGTCTCTGTTCCACAATTAATTTTGCAGATGTATA 657

QY 608 TGAGCCTGA 616
 Db 658 TCAGTCTCA 666

RESULT 12
 AA791818 385 bp mRNA linear EST 09-FEB-1998
 LOCUS v61c09.r1 Stratagene mouse skin (#937313) Mus musculus cDNA clone
 DEFINITION IMAGE:1150768 5', similar to SW:HK_HUMAN P51811 MEMBRANE TRANSPORT
 PROTEIN XK ;, mRNA sequence.
 ACCESSION AA791818
 VERSION AA791818.1 GI:2854773
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 385)
 AUTHORS Matra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.

DEFINITION Pan troglodytes DNA, clone: PTB-083B22.F, genomic survey sequence.
ACCESSION AG085155
VERSION AG085155.1 GI:16636957
KEYWORDS GSS.
SOURCE BAC Library clone: PTB-083B22.F.
ORGANISM Pan troglodytes
REFERENCE 1
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE BAC end sequences of Library PTB
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 691)
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: chimpbes@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/, Tel: 81-45-503-9111, Fax: 81-45-503-9170)
COMMENT Clones are derived from the chimpanzee BAC library PTB. This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI
FEATURES
Location/Qualifiers
1. .691
/organism="Pan troglodytes"
/db_xref="taxon:9598"
/clone="PTB-083B22.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"
230 a 151 c 127 g 181 t 2 others
BASE COUNT 230 a 151 c 127 g 181 t
ORIGIN
Query Match 8.1%; Score 113.2; DB 17; Length 691;
Best Local Similarity 60.9%; Pred. No. 7.7e-20;
Matches 184; Conservative 0; Mismatches 118; Indels 0; Gaps 0;
QY 640 AGAGTTGCTGCTAATGTTATTTCCCTGCTATCTGTACCTATGGGGCCACCCTTTGCAAT 699
DB 354 ATAGCATTGCTGATGACATGTTCCCTGTTATCAGTTACTTATGGGGCCATTCGCTGCAAT 295
QY 700 ATGTTGGCTATCCAGATCAAGTACGATGACTACAAAGATTTCGCTTGGGCCCACTAGAAGTC 759
DB 294 ATACTGGCCATCCAGATCAGCAATGATGATGATACCTACCCTAGCTACCGCGGATGATTC 235
QY 760 CTCTGCATCACCATCTCGGCGACATTTGAGATCACTTCCCGCCCTCTGATTTCTGGTGCTC 819
DB 234 TTCTGTGCTGATGTGGCGTTTTTGGAGGTTATCTCACGCTGTAGTGACTCTGGCATTT 175
QY 820 TTCTCAGCCATTTGAATTTGAAGGCTGTGCGCTTCCCTAGTCTCAACTTCTCATCATC 879
DB 174 TTCAATGCATCTCTGAAACTGGAAGGCTACCCCTTTTGTGTTAATCATATATTTTGTGTCA 115
QY 880 CTCTTTGAGCCCTGGATTAAGTTCTCGGAGAAGTGGTGGCCAGATGCCCAATAACATTTGAG 939
DB 114 TTGTTGCCACCGTGGCTGGAGTTTGGAAAGTGGAGCTCCGCGGACGTCTGATGTCGCG 55
QY 940 AA 941
DB 54 AA 53
RESULT 15

BI461092
LOCUS 603206970F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5272840 5',
DEFINITION mRNA sequence.
ACCESSION BI461092
VERSION BI461092.1 GI:15251748
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 666)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs@xmail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1688 row: k column: 17
High quality sequence stop: 663.
FEATURES
Location/Qualifiers
1. .666
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5272840"
/clone_lib="NIH_MGC_97"
/lab_host="DH10B"
/note="Organ: testis; Vector: pBluescriptR (modified pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."
125 a 212 c 197 g 132 t
BASE COUNT 125 a 212 c 197 g 132 t
ORIGIN
Query Match 7.9%; Score 109.2; DB 13; Length 666;
Best Local Similarity 53.7%; Pred. No. 9.5e-19;
Matches 278; Conservative 0; Mismatches 228; Indels 12; Gaps 2;
QY 157 TCCACCTTTTGTACTGTGGGAGGCTGCATCTGCTTTGTATCATGTTAGATCTATCGA 216
DB 156 TCCGTGTTCTCTGTTGTCGGCGAGACAACGGCGCTCAGCTGAGCAGCACTACCGC 215
QY 217 AAGAAATAGTGAACCTTACTGGATGACATACACCTTTCTCTTTATGTTTTCATCCATT 276
DB 216 TCGGGCGGGACCGCATGTGGCAGCGCTGACGTGCTTTTCTCGCTACTGCTTGGCGG 275
QY 277 ATGTCTCAGTTGACCCCTCATTTTGTCCACAGAGATCTAGCCAAAGATAAACCGCTATCA 336
DB 276 CTGTGGAGCTCAGCTTCTCTTGTACCCGACCTCAGCCGCGACCGCCCGCTCGTA 335
QY 337 TTATTTATGATCTAATCTCTCTTGGGACCTGTTATCAGATGTTTGGAGGCCATGATTAG 396
DB 336 CTGCTGTGACCTGCTGCAACTTGGGCCCTTTTCAGTGTGTTTGAAGTCTTCTGCATC 395
QY 397 TACCTCACTGTGGAAAGAAAGAGAGAGAGAGAGCCCTATGTGAGCCTCAACCCGAAG 456
DB 396 TACTTTC-----AGTCAGGCAACAATGAAGAGCCTTATGTGATATCACCAGAG 446
QY 457 A---AGATGCTATAGATGGGAGAGGTGCTGATAGATGGAGGTGGGCCACTCCATC 513
DB 447 AGGCAAAATGCCAAAAAATGGCCTCTCAGAGAGAGATTGAGAGAGGTGGGCCAGCAGAA 506

Qy	514	CGGACCTGGCTATGCACCGCAATGCCTACAAACGATATGTCACAGATCCAAGCTTCCTG	573
Db	507	GGCAAACTAATCACCACCGATCAGCGTTTCAGCCGGCGTCGGTGATCCAGGCTTTCCTG	566
Qy	574	GGCTCAGTGCCCCCAGCTGACCTATCAGCTCTATGTGAGCCTGATCTCTGCAGAGGTTCCC	633
Db	567	GGCTCAGCCCCCAGCTGACCTACAGCTGTACATAAGTGTCTATGCAGCAGGACGTCCT	626
Qy	634	CTGGTAGAGTTGTGCTAATGGTATTTTCCCTGGTATC	671
Db	627	GTGGAGAAGTCTCCTCATGACCATATCCCTGTTGTC	664

Search completed: March 10, 2003, 05:49:12
Job time : 1679.99 secs